

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2001-249698
(P2001-249698A)

(43)公開日 平成13年9月14日(2001.9.14)

(51) Int.Cl.⁷

識別記号

FI

テーマコード・(参考)

G 1 0 L 19/02

13/00

H03M 7/30

H 0 4 B 14/04

H O 3 M 7/30

H04B 14/04

G 1 0 L 7/04

9/00

Z 5 J 0 6 4

Z 5K041

G 9A001

L

審査請求 有 請求項の数6 O.L (全 20 頁)

(21)出願番号

特願2000-60932(P2000-60932)

(22) 出題日

平成12年3月6日(2000.3.6)

(71)出願人 399004577

株式会社ワイ・アール・ビー高機能移動体
通信研究所

神奈川県横須賀市光の丘3番2号

(72) 發明者 蘆 照夫

神奈川県横須賀市光の丘3番2号 株式会社ワイ・アール・ピー高機能移動体通信研究所内

(74) 代理人 100106459

弁理士 高橋 英生 (外3名)

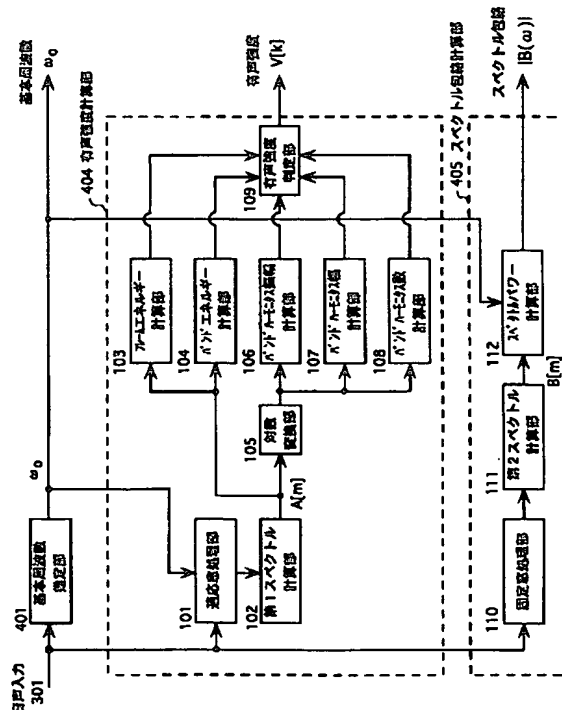
[最終頁に続く](#)

(54)【発明の名称】 音声符号化パラメータ取得方法、音声復号方法及び装置

(57) 【要約】

【課題】 周波数スペクトル雑音に強い有声／無声判定を行う。計算量の少ない無声音声の合成を行う。

【解決手段】 適応窓処理部 101 で音声基本周波数に応じて窓長を適応的に調整し、スペクトル計算部 102 で音声セグメントのスペクトルを求める。フレームエネルギー計算部 103 で音声セグメントのスペクトルパワー、バンドエネルギー計算部 104 で周波数スペクトル帯域を分割した各周波数バンドのスペクトルパワー、バンドハーモニクス振幅計算部 106 とバンドハーモニクス幅計算部 107 で各ハーモニクスの強度と幅、バンドハーモニクス数計算部 108 でハーモニクスの数を算出し、それらに基づいて、有声強度判定部 109 で各周波数バンドの有声強度 $V[k]$ を算出する。音声復号部では、対称ランダム系列と反対称ランダム系列から、無声に対応する帯域に相当する区間を取りだし、逆FFTして時間領域の信号に変換した後、線形補間してフレーム長の無声音声を得る。



1

【特許請求の範囲】

【請求項 1】 デジタル化された音声信号を、ある一定の繰り返し周期で、所定のセグメント長で抜き取った音声セグメントから音声符号化パラメータを取得する音声符号化パラメータの取得方法であって、

前記音声セグメントから音声基本周波数を取得するステップ、

前記音声基本周波数により決定される可変長の適応窓により前記音声信号を抜き取った可変長セグメントから第 1 の周波数スペクトルを取得するステップ、

前記音声信号を固定長の窓により抜き取った固定長セグメントから第 2 の周波数スペクトルを取得するステップ、

前記第 1 の周波数スペクトルを複数の周波数バンドに分割するステップ、

前記第 1 の周波数スペクトルの周波数スペクトルパワー、前記各周波数バンドの周波数スペクトルパワー、前記各周波数バンドに含まれるハーモニクス数、各ハーモニクスのハーモニクス振幅およびハーモニクス帯域幅により前記各周波数バンド毎の有声強度を決定するステップ、および、

前記第 2 の周波数スペクトルから前記音声基本周波数の整数倍の周波数を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーを計算するステップを含むことを特徴とする音声符号化パラメータの取得方法。

【請求項 2】 前記可変長の適応窓の長さは、前記可変長の適応窓の周波数スペクトル分布の帯域幅と前記音声基本周波数の関係により決定される事の特徴とする前記請求項 1 記載の音声符号化パラメータの取得方法。

【請求項 3】 前記可変長の適応窓は、前記音声基本周波数に対応する周期の 4 倍以上の長さを持つハミング窓である事の特徴とする前記請求項 1 記載の音声符号化パラメータの取得方法。

【請求項 4】 デジタル化された音声信号を、ある一定の繰り返し周期で抜き取った音声セグメントの音声基本周波数と、該音声セグメントの周波数スペクトルを音声基本周波数の整数倍を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーと、前記音声セグメントの周波数スペクトルを複数の周波数バンドに分割した各周波数バンドが有声音か無声音かを判別した判別情報からなる音声符号化パラメータによって音声合成する音声復号方法であって、

前記判別情報が有声を示す前記周波数バンドでは、その中心周波数が前記音声基本周波数の整数倍の周波数を持ち、且つ、対応する前記ハーモニクス帯域のスペクトルパワーと同等になる振幅を持った正弦波群を生成し、また、前記判別情報が無声を示す周波数バンドでは、中心対称ランダム系列と中心反対称ランダム系列を雑音信号

2

の周波数スペクトル系列の実部と虚部と見なし、前記 2 つのランダム系列から該周波数バンドに対応する区間を抽出し、対応する前記ハーモニクス帯域のスペクトルパワーと同じになる様に振幅調整した後、逆フーリエ変換によりその実部を得て無声フレーム信号とし、1 つ前のセグメントの無声フレーム信号と今回得た前記無声フレーム信号間で線形補間することにより無声音声を生成した後、前記生成した正弦波群と加算して合成音声を得る音声復号方法。

10 【請求項 5】 デジタル化された音声信号を、ある一定の繰り返し周期で、所定のセグメント長で抜き取った音声セグメントから音声符号化パラメータを取得する音声符号化パラメータの取得装置であって、

前記音声セグメントから音声基本周波数を取得する手段、

前記音声基本周波数により決定される可変長の適応窓により前記音声信号を抜き取った可変長セグメントにより第 1 の周波数スペクトルを取得する手段、

20 前記音声信号を固定長の窓により抜き取った固定長セグメントにより第 2 の周波数スペクトルを取得する手段、前記第 1 の周波数スペクトルを複数の周波数バンドに分割する手段、

前記第 1 の周波数スペクトルから周波数スペクトルパワー、前記各周波数バンドの周波数スペクトルパワー、前記各周波数バンドに含まれるハーモニクス数、各ハーモニクスのハーモニクス振幅およびハーモニクス帯域幅により前記各周波数バンド毎の有声強度を決定する手段、および、

30 前記第 2 の周波数スペクトルから前記音声基本周波数の整数倍の周波数を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーを計算する手段を有することを特徴とする音声符号化パラメータの取得装置。

【請求項 6】 デジタル化された音声信号を、ある一定の繰り返し周期で抜き取った音声セグメントの音声基本周波数と、該音声セグメントの周波数スペクトルを音声基本周波数の整数倍を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーと、前記音声セグメントの周波数スペクトルを複数の周波数バンドに分割した各周波数バンドが有声音か無声音かを判別した判別情報からなる音声符号化パラメータによって音声合成する音声復号装置であって、

前記判別情報が有声を示す前記周波数バンドでは、その中心周波数が前記音声基本周波数の整数倍の周波数を持ち、且つ、対応する前記ハーモニクス帯域のスペクトルパワーと同等になる振幅を持った正弦波群を生成する手段、

50 中心対称ランダム系列と中心反対称ランダム系列の雑音信号を発生する手段、

3

前記 2 つのランダム系列から前記判別情報が無声を示す前記周波数バンドに対応する区間を抽出する手段、抽出したランダム系列の雑音信号を、そのスペクトルパワーが前記判別情報が無声を示す前記周波数バンドに対応するハーモニクス帯域のスペクトルパワーと同じになる様に振幅調整する手段、該振幅調整されたランダム系列の雑音信号を逆フーリエ変換し、無声フレーム信号を生成する手段、1 つ前のセグメントの無声フレーム信号と今回の無声フレーム信号を線形補間することにより無声音声を生成する手段、および、前記生成された正弦波群と生成された無声音声を加算する手段を有することを特徴とする音声復号装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声信号をデジタル化して所定の時間間隔毎にその特徴を表す音声符号化パラメータを取得する音声符号化パラメータ取得方法および装置、ならびに、前記音声符号化パラメータに基づいて元の音声信号を合成する音声復号方法及び装置に関するものであり、音声符号化パラメータを符号化して伝送または蓄積し、伝送先または蓄積先から必要な時に音声符号化パラメータを復元し、復元した音声符号化パラメータから音声信号を合成して音声伝えるデジタル携帯電話やデジタル音声蓄積装置などに使用して好適なものである。

【0002】

【従来の技術】デジタル化された音声信号は、データ圧縮、誤り処理、多重化などさまざまなデジタル信号処理が可能になるため、固定電話や携帯電話に限らず音声を利用するマルチメディアシステムなどに広く取り入れられている。アナログの音声信号をデジタル化するには、一般に入力音声周波数帯域の 2 倍以上の標本化周波数で標本化し、耳に識別できない程度の量子化ステップで量子化することが必要のため、アナログ信号と比較し広い伝送周波数帯域幅を必要とする。そのため、一旦デジタル化された音声信号は、要求される音声品質に応じてさまざまな符号化方式や変調方式によりデータの圧縮が行われている。音声の持つ特徴を積極的に利用する事により、効率的な圧縮を行う事が出来る。例えば、適応型差分パルス符号変調 (ADPCM) 方式は音波形の周期性や人間の聴覚感度の対数特性を利用した波形符号化方式で、12.8 kbps のデジタル音声を 3.2 kbps 程度に圧縮して圧縮前と変わらない音声品質を得ており、電話の基幹伝送や PHS システムに利用されている。波形符号化方式は標本化点を最低 1 ビットで表現するものであるため、標本化周波数が 8 KHz の場合では原理的に 8 kbps 以下に音声符号化速度を下げる事は出来ない。

【0003】低い音声符号化速度を得るために、音声を所定の時間間隔のセグメントに分割し、そのセグメント

4

毎に、音声合成パラメータと残差音源信号を伝送する符号励振線形予測 (CELP) を基本とする方式がある。日本の携帯無線電話で用いられている VSELP や PSI-CELP 方式は、20 msec や 40 msec 間隔の音声信号の線形予測分析により得られる人間の声道フィルタ特性を近似する線形予測係数 (LPC) と、聴感的に入力音声に近い波形を合成する事が出来る残差音源信号を符号化する事で低い音声符号化速度を実現している。また残差音源信号を効率良く符号化するために、複数の残差音源波形を持った符号帳を用意し、その符号帳のエントリ番号と利得を伝送している。これらの詳細は電波産業会の規格書 RCR-STD 27 F に詳しく記載されている。この CELP を基本とする方式は、適切な大きさの符号帳をうまく設計する事で音声符号化速度 3 ~ 4 kbps 程度まで実現されている。

【0004】更に低い音声符号化速度を得るために、音声合成パラメータのみを伝送して上記の CELP 方式における音源符号帳を用いない音声符号化を行う方式がある。米国国防省の標準音声符号化方式の FS-1015 は、ピッチ周波数、LPC 係数、ルート二乗平均振幅、有声/無声判定情報の音声合成パラメータにより音声符号化・復号を行う LPC ボコーダ (Vocoder) 方式の音声符号化方式で、2.4 kbps の音声符号化速度を得ている。この方式は音声の特徴を積極的に利用しているが、合成音声的な音質になり、特に背景雑音下で復号音声品質が著しく劣化する欠点を有していた。また、衛星携帯電話に一部使用されている IMBE (Improved Multiband Excitation) 方式は、音声時間セグメントを周波数領域に変換して音声ピッチ、音声ハーモニクス振幅、周波数帯域を複数に分割した周波数バンドの有声/非有声情報で音声符号化を行う方式で、音声セグメントの各バンド毎に有声音モデルと無声音モデルを選択して合成するために、背景雑音下や混合音声の場合にも合成音声の劣化が少なく、前記 LPC ボコーダに比べて優れていると報告されている。

【0005】図 14 は、一般的な音声符号化伝送装置の構成を示した図である。音声符号化パラメータ抽出部 302 は音声入力端子 301 から入力された標本化・量子化された音声デジタル信号を、所定の時間間隔のセグメントに分割し、そのセグメント毎に音声符号化パラメータを抽出する。抽出する音声符号化パラメータは音声符号化方式により決定され、例えば前記の IMBE 方式では、音声ピッチ、音声ハーモニクスの振幅、各周波数バンドの有声/無声情報である。パラメータ符号化部 303 は、抽出した音声符号化パラメータを効果的に符号化して符号量を低減せしめ、送信部 304 を介して伝送路 305 に送り出す。パラメータ復号部 307 は受信部 306 で受け取った符号を復号して音声符号化パラメータを復元し、音声合成部 308 は音声符号化パラメータ抽出部の動作と逆の動作により合成音声を作成し音声を

5

力端子 309 から音声デジタル信号を出力する。

【0006】図 15 は前記 IMBE 方式の場合における前記音声符号化パラメータ抽出部 302 の内部構成図である。デジタル入力音声信号 301 は基本周波数推定部 401 に入力され、ここで音声の基本周波数が推定される。基本周波数の推定には、自己相関関数や周波数スペクトルの対数の逆フーリエ変換であるケプストラムのピークを検出する方法など多くの方法があり、例えば、古井著「デジタル音声処理」東海大学出版会、1985 年 9 月 25 日 等に記載されている。周波数スペクトル計算部 402 では、ハミング窓等の窓関数により切り出した有限長の音声セグメントを周波数分析して音声周波数スペクトルを得る。基本周波数修正部 403 は、前記基本周波数推定部 401 で推定された音声基本周波数の近傍の周波数範囲で A-b-S (Analysis-by-Synthesis) 手法により合成スペクトルと前記周波数スペクトル計算部 402 で算出した音声周波数スペクトルとの誤差最小条件により修正した基本周波数 ω_0 を得る。有声強度計算部 404 は該修正された基本周波数 ω_0 に基づいて、周波数帯域を複数の周波数バンド（周波数区間） k ($k=1, 2, \dots, K$) に分割し、各周波数バンド毎に合成された合成スペクトルと音声周波数スペクトルの誤差を計算し、閾値判定により有声/無声情報 $V[k]$ を出力する。スペクトル包絡計算部 405 は有声/無声情報 $V[k]$ により、有声バンドでは A-b-S 手法で求めた各ハーモニクスの振幅、無声バンドでは各ハーモニクスの持つ周波数帯域での周波数スペクトルのルート二乗平均値（RMS 値）をスペクトル包絡絶対値 $|A(\omega)|$ として出力する。

【0007】図 16 は、前記 IMBE 方式の場合における前記音声合成部 308 の内部構成を示す図である。この図に示すように、音声合成部 308 は、有声音声合成部 508 と無声音声合成部 509 に大別される。有声音声合成部 508 において、高調波音源部 501 では、有声/無声情報 $V[k]$ と基本周波数 ω_0 により、有声と判定される周波数区間において、基本周波数 ω_0 とその高調波の音源をスペクトル包絡 $|A(\omega)|$ に対応する振幅で駆動して複数の音源信号を生成する。高調波加算部 502 では高調波音源部 501 で発生した複数の音源信号を加算合成し、有声バンドに対応する音声信号を生成する。また、無声音声合成部 509 において、雑音音源部 503 は、白色雑音を生成し、周波数変換部 504 で適当な窓関数により処理した後、フーリエ変換（FFT）して周波数領域の信号に変換する。雑音抽出部 505 で、周波数領域の信号に変換された白色雑音から $V[k]$ により無声と指定された周波数バンドの白色雑音スペクトルを取りだし、スペクトル包絡 $|A(\omega)|$ の周波数バンド毎のパワーに一致する様に各スペクトルの振幅を調整する。

6

逆周波数変換部 506 では無声バンドに対応する雑音区間の周波数スペクトルを逆フーリエ変換（IFFT）することにより音声波形に変換する。加算部 507 では、有声音声合成部 508 の高調波加算部 502 からの有声音声波形と、無声音声合成部 509 の逆周波数変換部 506 で時間軸波形信号に変換した無声音声波形とを加算し、最終的な有声音と無声音を持った合成音声を得ている。この IMBE 方式の詳細は、“Multiband Excitation Vocoder”, IEEE Transactions on Acoustics, speech, and signal processing, vol. 36, No. 8, August 1988, pp1223-1235 に詳しく記載されている。このように、音声をデジタル化して低ビットレートの音声符号化を実現する方法として、音声合成モデルに基づく音声符号化パラメータを抽出して符号化を行う IMBE 方式等の分析合成型の音声符号化方式が提案されている。

【0008】

【発明が解決しようとする課題】以上述べた様に、低ビットレート音声符号化のためには分析合成型の音声符号化方式が有効であるが、残差音源信号を用いず音声合成パラメータでのみ音声の合成を行うため、符号化方式によっては合成音的な音質になりやすい。低ビットレート音声符号化を実現する分析合成型の音声符号化としての IMBE 方式は、入力音声をセグメントに分割して音声フレームを切り出し、フレームの周波数帯域を複数のバンドに分割してそのバンドに含まれる周波数成分が有声か無声かを判定し、バンド毎に有声音合成モデルと無声音合成モデルを設定し、それらを加算して合成音声を得る事で、背景雑音等が入った有声音や、有声と無声の混在したフレームでの合成音声品質を改善している。符号化パラメータとしての、音声基本周波数（音声ピッチ）、有声/無声情報、スペクトル包絡情報の推定の正確さは、再生音声の品質を決める上で重要である。音声基本周波数は前述の自己相関法により求める事も出来るが、上記 IMBE 方式では整数倍ピッチまで拡張した自己相関関数で評価を行い $1/2$ ピッチ精度で基本周波数推定を行う方法が示されている。また、スペクトル包絡を求めるには、抽出された基本周波数 ω_0 と周波数分析窓の周波数スペクトルを用いて前述の A-b-S 手法で推定されるが、前記推定された基本周波数の精度ではスペクトル包絡の推定には精度が不足するため、推定基本周波数の近傍を $1/4$ ピッチ精度で探索しながらスペクトル包絡を同時に推定する手法を取っている。

【0009】このスペクトル包絡を求める手順は以下のとおりである。まず、入力音声セグメントの信号 $s(n)$ を周波数分析窓 $w_R(n)$ で範囲を -110 から 110 サンプルに制限した後、256 段の FFT により周波数スペクトル $S_w(m)$ を式 (1) により得る。

【数 1】

$$S_w(m) = \sum_{n=-110}^{110} s(n)w_R(n)e^{-j2\pi mn/256} \quad \text{for } -128 \leq m \leq 127 \quad (1)$$

次に、基本周波数 ω_0 の L 次高調波 ($L=1, 2, \dots, L_{\max}$; $(L_{\max}+0.5) \cdot \omega_0 < 2\pi$) を中心周波数として上記周波数分析窓 $w_R(n)$ の周波数スペクトル $E_w(\omega)$ の広がりを持ち、式 (2) で示す包絡値 A_L のスペクトルの総和

$$S_w(m) \cong \hat{S}_w(m, \omega_0) = \sum_{L=1}^{L_{\max}} A_L \cdot E_w(\omega - L\omega_0) \quad \text{但し } (L_{\max}+0.5) \cdot \omega_0 \leq 2\pi \quad (2)$$

【0010】この時、 ω_0 を $1/4$ ピッチ精度で変化させながら、誤差最小二乗法で各高調波の包絡値 A_L を算出し、求めた ω_0 と A_L によりスペクトル誤差評価値 $E(\omega_0)$

【数3】

$$E(\omega_0) = \sum_m |S_w(m) - \hat{S}_w(m, \omega_0)|^2 \quad (3)$$

$$D_k = \sum_{m=a1}^{M-1} |S_w(m) - \hat{S}_w(m, \omega_0)|^2 / \sum_{m=a1}^{M-1} |S_w(m)|^2 \quad (4)$$

【0011】ここで、 ω_0 を探索する場合の変化ステップと、実際の音声基本周波数の誤差が判定結果に及ぼす影響について考察する。音声基本周波数は、個人や男女により開きはあるが、男性では中心が約 125 Hz、女性は約 2 倍の基本周波数を持ち、全体で 70 Hz から 400 Hz の範囲にある。また評価する基本周波数の誤差は L 次高調波では L 倍の周波数誤差に拡大される。表 1 は音声基本周波数 $\omega_0 (=2\pi f_0)$ のピッチ誤差 ΔP により発生する周波数誤差 Δf と 2 kHz 付近の高調波領域での周

で上記音声セグメントのスペクトル $S_w(m)$ を近似し、その個々の高調波の包絡値 A_L を求める。

【数2】

が最小となる ω_0 を基本周波数修正値とし、その時の各高調波の振幅 A_L をスペクトル包絡値とする。また、

(4) 式に示すスペクトル誤差相対値 D_k を閾値判定する事により各周波数帯域 ($a1 \sim b1$) の有声強度を推定している。

【数4】

波数誤差 Δf (2 kHz) を式 (5) により計算したものである。ここで f_s は音声セグメントの標準化周波数である。

【数5】

$$\Delta f = f_0 - f_s / (P + \Delta P)$$

(5)

$$\Delta f_{2kHz} = \Delta f \cdot 2000 / f_0$$

【0012】

【表1】

基本周波数 f_0 (Hz)	ピッチ $P = f_s / f_0$	ピッチ誤差 ΔP (\pm ピッチ)	周波数誤差 Δf (\pm Hz)	
			$f = f_0$	$f \approx 2\text{kHz}$
100	80	1	1.2	25
		0.5	0.6	12
		0.25	0.3	6
200	40	1	5.0	50
		0.5	2.5	25
		0.25	1.2	12
300	27	1	11.2	75
		0.5	5.6	38
		0.25	2.8	19

表1 ピッチ誤差による高調波周波数誤差 (標準化周波数 $f_s=8000\text{Hz}$)

【0013】表1からわかるように実際の音声基本周波数 (音声ピッチ) の推定誤差 ΔP が1ピッチの場合では、2 kHz 付近の高調波の周波数誤差は $\pm 25 \sim \pm 75\text{Hz}$ まで及び、256段FFTで周波数分析した場合のスペクトル間隔 $8000/256=31.25\text{Hz}$ 以上になる。また、 $f_0=300\text{Hz}$ では、 ΔP が0.5ピッチの時に2 kHz 付近の高調波スペクトル誤差は38Hzになり、 ΔP が0.25の場合によりやく19Hzになる。一方、周波数分析窓をハミング窓と

して式 (4) で計算したFFT256段の場合のスペクトル評価誤差を図17に示す。例えば、 $f_0=275\text{Hz}$ の基本周波数を持つ標準的な女性の場合、基本周波数ピッチは $P_i=29$ であるが、基本周波数の推定を $P_i=28$ と推定し、-1ピッチの誤差があった場合、推定した基本周波数は $f_0=8000/(29-1)=285.7\text{Hz}$ 、基本周波数誤差は $\Delta f_0=10.7\text{Hz}$ になり、図17より正規化スペクトル誤差 $D_k=0.1$ になる。更に、推定ピッチ $P_i=27$ の

場合では、 -2 ピッチの推定誤差、基本周波数誤差は 21 Hz になり、 $Dk=0.3$ になり、正規化スペクトル誤差による有声/無声判定へ与える影響が大きい。さらに、ハーモニクス高調波の場合には基本周波数誤差は高調波次数分拡大される。例えば $1/4$ ピッチ誤差の場合では、推定基本周波数は $8000/(29-0.25)=278.26\text{ (Hz)}$ になり、基本周波数誤差 $\Delta f_0=3.26\text{ Hz}$ であるが、 2 kHz 付近では $2000/275$ 倍拡大して 23.7 Hz の周波数誤差になり、図 17 より正規化スペクトル誤差は 0.01 程度から 0.35 以上にまで拡大し、有声/無声判定誤りの原因となる。有声/無声情報やスペクトル包絡情報は音声セグメント全体を特徴づけるパラメータでもあり、これらの推定の誤りは、既に述べた様に符号化音声の品質に与える影響が大きい。

【0014】また、IMBE方式における音声復号においては、前記図 16 に示したように、無声音ではランダム雑音音源を周波数変換 (FFT) し、音声符号化パラメータにより指示された無声音の周波数範囲のみ抽出した後、逆周波数変換 (IFFT) して無声音の音声合成している。この場合、周波数変換が 2 段必要であり、特に、符号化音声品質を高めるため音声セグメントの更新周期を短く設定した場合に演算負荷が大きいという欠点がある。

【0015】そこで、本発明は、音声基本周波数の変化に係らず、高精度の有声強度判定を行うことができ、スペクトル雑音に対し誤り耐性の強い音声符号化方法および装置を提供することを目的としている。また、演算負荷の小さな音声復号方法および装置を提供することを目的としている。

【0016】

【課題を解決するための手段】上記目的を達成するために、本発明の音声符号化パラメータの取得方法は、デジタル化された音声信号を、ある一定の繰り返し周期で、所定のセグメント長で抜き取った音声セグメントから音声符号化パラメータを取得する音声符号化パラメータの取得方法であって、前記音声セグメントから音声基本周波数取得するステップ、前記音声基本周波数により決定される可変長の適応窓により前記音声信号を抜き取った可変長セグメントから第 1 の周波数スペクトルを取得するステップ、前記音声信号を固定長の窓により抜き取った固定長セグメントから第 2 の周波数スペクトルを取得するステップ、前記第 1 の周波数スペクトルを複数の周波数バンドに分割するステップ、前記第 1 の周波数スペクトルの周波数スペクトルパワー、前記各周波数バンドの周波数スペクトルパワー、前記各周波数バンドに含まれるハーモニクス数、各ハーモニクスのハーモニクス振幅およびハーモニクス帯域幅により前記各周波数バンド毎の有声強度を決定するステップ、および、前記第 2 の周波数スペクトルから前記音声基本周波数の整数倍の周波数を中心としてその周波数帯域幅が音声基本周波数

になる様に分割した各ハーモニクス帯域のスペクトルパワーを計算するステップを含むものである。また、前記可変長の適応窓の長さは、前記可変長の適応窓の周波数スペクトル分布の帯域幅と前記音声基本周波数の関係により決定されるものである。さらに、前記可変長の適応窓は、前記音声基本周波数に対応する周期の 4 倍以上の長さを持つハミング窓とされている。

【0017】さらにまた、本発明の音声復号方法は、デジタル化された音声信号を、ある一定の繰り返し周期で抜き取った音声セグメントの音声基本周波数と、該音声セグメントの周波数スペクトルを音声基本周波数の整数倍を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーと、前記音声セグメントの周波数スペクトルを複数の周波数バンドに分割した各周波数バンドが有声音か無声音かを判別した判別情報からなる音声符号化パラメータによって音声合成する音声復号方法であって、前記判別情報が有声を示す前記周波数バンドでは、その中心周波数が前記音声基本周波数の整数倍の周波数を持ち、且つ、対応する前記ハーモニクス帯域のスペクトルパワーと同等になる振幅を持った正弦波群を生成し、また、前記判別情報が無声を示す周波数バンドでは、中心対称ランダム系列と中心反対称ランダム系列を雑音信号の周波数スペクトル系列の実部と虚部と見なし、前記 2 つのランダム系列から該周波数バンドに対応する区間を抽出し、対応する前記ハーモニクス帯域のスペクトルパワーと同じになる様に振幅調整した後、逆フーリエ変換によりその実部を得て無声フレーム信号とし、1 つ前のセグメントの無声フレーム信号と今回得た前記無声フレーム信号間で線形補間することにより無声音声を生成した後、前記生成した正弦波群と加算して合成音声を得るものである。

【0018】さらにまた、本発明の音声符号化パラメータの取得装置は、デジタル化された音声信号を、ある一定の繰り返し周期で、所定のセグメント長で抜き取った音声セグメントから音声符号化パラメータを取得する音声符号化パラメータの取得装置であって、前記音声セグメントから音声基本周波数取得する手段、前記音声基本周波数により決定される可変長の適応窓により前記音声信号を抜き取った可変長セグメントにより第 1 の周波数スペクトルを取得する手段、前記音声信号を固定長の窓により抜き取った固定長セグメントにより第 2 の周波数スペクトルを取得する手段、前記第 1 の周波数スペクトルを複数の周波数バンドに分割する手段、前記第 1 の周波数スペクトルから周波数スペクトルパワー、前記各周波数バンドの周波数スペクトルパワー、前記各周波数バンドに含まれるハーモニクス数、各ハーモニクスのハーモニクス振幅およびハーモニクス帯域幅により前記各周波数バンド毎の有声強度を決定する手段、および、前記第 2 の周波数スペクトルから前記音声基本周波数の整

数倍の周波数を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーを計算する手段を有するものである。

【0019】さらにまた、本発明の音声復号装置は、デジタル化された音声信号を、ある一定の繰り返し周期で抜き取った音声セグメントの音声基本周波数と、該音声セグメントの周波数スペクトルを音声基本周波数の整数倍を中心としてその周波数帯域幅が音声基本周波数になる様に分割した各ハーモニクス帯域のスペクトルパワーと、前記音声セグメントの周波数スペクトルを複数の周波数バンドに分割した各周波数バンドが有声音か無声音かを判別した判別情報からなる音声符号化パラメータによって音声合成する音声復号装置であって、前記判別情報が有声を示す前記周波数バンドでは、その中心周波数が前記音声基本周波数の整数倍の周波数を持ち、且つ、対応する前記ハーモニクス帯域のスペクトルパワーと同等になる振幅を持った正弦波群を生成する手段、中心対称ランダム系列と中心反対称ランダム系列の雑音信号を発生する手段、前記2つのランダム系列から前記判別情報が無声を示す前記周波数バンドに対応する区間を抽出する手段、抽出したランダム系列の雑音信号を、そのスペクトルパワーが前記判別情報が無声を示す前記周波数バンドに対応するハーモニクス帯域のスペクトルパワーと同じになる様に振幅調整する手段、該振幅調整されたランダム系列の雑音信号を逆フーリエ変換し、無声フレーム信号を生成する手段、1つ前のセグメントの無声フレーム信号と今回の無声フレーム信号を線形補間することにより無声音声を生成する手段、および、前記生成された正弦波群と生成された無声音声を加算する手段を有するものである。

【0020】

【発明の実施の形態】本発明の音声符号化パラメータの取得方法、音声復号方法および装置は、例えば、音声符号化、特に低ビットレートの音声符号化での音声符号化パラメータを安定に推定する方法および装置、さらには推定した音声符号化パラメータによって音声復号する方法および装置に組み込み使用することができるが、ここでは、前記図14に示した音声符号化伝送装置の音声符号化パラメータ抽出部302、および音声合成部308に本発明を適用した場合を例にとって説明する。また、本発明は、種々の音声符号化方式に適用することが可能であるが、ここでは、IMBE方式に適用した場合を例にとって説明する。

【0021】図1は本発明の音声符号化パラメータの取得方法が適用された音声符号化パラメータ抽出部のブロック構成図である。この図に示すように、本発明の音声符号化パラメータ抽出部は、入力音声信号301からその基本周波数 ω_0 を推定する基本周波数推定部401、入力音声信号フレームを周波数分析して得た周波数スペクトルを複数の周波数バンドに分割し、各バンドごとに

その有聲/無声を示す有聲強度情報 $V[k]$ を出力する有聲強度計算部404、および、入力音声信号を固定長の窓を用いて周波数分析し、スペクトル包絡 $|B(\omega)|$ を計算するスペクトル包絡計算部405の3つの部分から構成されている。

【0022】ここで、本発明の音声符号化パラメータの取得方法においては、従来方式のように合成音声と入力音声の周波数スペクトル誤差を有聲・無聲の評価値とすることはせず、入力音声の周波数スペクトルのある周波数バンドに含まれる音声のハーモニクス振幅を入力音声スペクトル振幅から計測して、そのハーモニクス振幅を有聲強度あるいは有聲/無聲の判定の評価値としている。そして、前記入力音声の周波数スペクトルを計測するにあたり、スペクトル分析窓の幅を入力音声の推定基本周波数に応じて適応的に調節する事で周波数分解能を調節し、むやみに時間分解能を低下する事なく、必要な周波数分解能を得る手法を取っている。また、各周波数バンドに含まれるハーモニクスの数を計測し、そのハーモニクス数をもう一つの評価値として、期待されるハーモニクス数にどれだけ近いかを判定する。更に、各ハーモニクスの周波数の幅（ハーモニクス幅）を計測して前記スペクトル分析窓により期待されるハーモニクス幅にどれだけ近いかを判定することにより、判定の確実性を向上させている。さらにまた、入力音声のパワー（エネルギー）が小さい場合は無声であるとの知見から、入力音声の周波数スペクトルパワーさらには各周波数バンドの音声周波数スペクトルパワーも評価値に加えるようにしている。また、スペクトル包絡の抽出にあたっては、前記スペクトル分析窓と分析長の異なる固定長の窓を用いた第2の周波数分析により入力音声の周波数スペクトルを取りだし、推定音声基本周波数の整数倍の周波数間隔毎に音声基本周波数幅の領域にあるスペクトルパワーの平方根として抽出している。

【0023】この理由について、図3、図4、図5を用いてさらに説明する。図3はほとんど有聲音声で出来ている音声セグメントの周波数スペクトル振幅値（対数値）の例である。横軸は256点の高速離散フーリエ変換（FFT）した場合の離散周波数である。この図に示すように、スペクトル振幅にはある一定の間隔で適度の幅を持った明瞭な高調波スペクトルが観測されており、その対数振幅や幅も広範囲の周波数にわたり安定な振幅を持っている。この事から、ある周波数バンド内のハーモニクス振幅とその数は、基本周波数 ω_0 の推定誤差の影響を受けずに計測できる事が予想できる。また、図4は無聲音声が多い音声セグメントの周波数スペクトル振幅値（対数値）の例である。この場合は、定められた周波数バンド内でのハーモニクス振幅やハーモニクス幅は小さく、また一定レベル以上のハーモニクスの数も少なくなっている事が読みとれ、その値は基本周波数 ω_0 の推定誤差 $\Delta\omega_0$ の影響をあまり受けない事も読み取れ

る。以上の考察により、有声／無声の判定に、周波数スペクトル振幅対数値から計測したハーモニクス振幅、ある閾値以上の振幅を持った有効なハーモニクス数、ハーモニクスの幅、さらには入力音声のパワー、周波数バンドの音声パワーを判定評価に使用するようにしている。

【0024】また、周波数分析窓の長さ（時間範囲） T_w (sec)又は L_w (サンプル)と音声基本周波数 f_0 (Hz)又はピッチ P (サンプル)の関係を考察すると、図5の様にスペクトル振幅には基本周波数の整数倍にハーモニクス中心が現れ、周波数分析窓がハミング窓の場合には各ハーモニクスの帯域幅は $4/T_w$ になる。従って、ハーモニクスの谷が隣のハーモニクスの谷より中心に侵入しない事を条件として、式(6)により周波数分析窓の長さを決める。

【数6】

$$2/T_w \leq f_0/2$$

または

$$L_w \geq 4P$$

(6)

【0025】なお、周波数分析窓の長さは、式(6)に

$$w_s(n) = 0.54 + 0.46 \cos\left(\frac{2\pi \cdot (n - \frac{M-1}{2})(M/L_w)}{M-1}\right); \quad (-\pi \leq \phi \leq \pi) \quad (8)$$

$$= 0; \quad (|\phi| > \pi)$$

ここで、 $n = 0, 1, 2, \dots, M-1$, M はFFT段数

【0026】さらに、本発明では、基本周波数 f_0 の間、に何本のスペクトル本数 p を設定するかで必要なFFT段数の M を決める事が出来る。例えば標準化周波数が8

$$8000/(f_0/p) \cong M$$

ここで、最小の基本周波数を60Hzとした場合では、 $p = 4$ の場合には $M = 533$ になり、512段程度のFFT段数が必要である事がわかる。

【0027】以上のように、本発明の音声符号化パラメータ取得方法によれば、音声基本周波数（ピッチ）により適応的に窓サイズが設定される周波数分析窓を用い、基本周波数の範囲から決定される段数のFFTにより周波数スペクトルを得ているため、ハーモニクス間のスペクトル相互干渉を少なくすることができる。そして、このようにして得た周波数スペクトルから、ハーモニクス振幅、ハーモニクス数、ハーモニクス幅、フレームエネルギー（フレームパワー）、バンドエネルギー（バンドパワー）を計測し、各周波数バンド毎の有声／無声情報を得るようにしているため、有声／無声の判定に詳細ピッチが不要となり、ピッチ誤りに起因する判定誤りの可能性を減少させることができる。また、スペクトル包絡情報の取得は、前記有声／無声判定とは分離して行うようにし、固定窓サイズのFFTによりハーモニクス帯域毎のエネルギーの平方根から得るようにしている。したがって、無声／有声判定の誤りがあってもそれに

より基本ピッチの4倍として基本周波数に比例して変化させても良いが、実用的には基本ピッチにより何段階かに分類して設定しても良い。例えば、ピッチが20増加する毎に切り替えて、式(7)により第1の周波数分析窓をピッチの変化に応じて設定しても良い。

【数7】

$$L_w = \text{ceil}(P/20) \cdot 80 + 1$$

$$\text{if } (L_w < 161) \text{ then } L_w = 161$$

(7)

$$\text{if } (L_w < 481) \text{ then } L_w = 481$$

ここで $\text{ceil}(x)$ は x を超える最小の整数を与える関数であり、また、分析窓長は中心対称の奇数で且つ分析窓長がピッチ範囲より想定される長さ以外になる事を防止するため制限をしている。図6に、抽出した基本ピッチ P に対する L_w の設定例を示す。また、窓関数として式(8)に示す適応ハミング窓を用いた場合の窓関数値 $w_s(n)$ の計算結果を図7に示す。（ただしFFT段数 $M = 512$ として計算した）

【数8】

kHzの場合には式(9)により決定できる。

【数9】

$$(9)$$

影響を受けないスペクトル包絡情報を得ることができる。

【0028】図2は、本発明の音声復号方法が適用された音声復号装置の一構成例を示すブロック図である。この図に示すように、音声復号装置は、復号された音声パラメータのうちの基本周波数 ω_0 とスペクトル包絡 $|B(\omega)|$ が入力され有声音声合成部508、復号された音声符号化パラメータのうちの有声／無声情報（有声強度情報） $V[k]$ および前記スペクトル包絡 $|B(\omega)|$ が入力される無声音声合成部509、および、加算部507から構成されている。ここで、有声音声合成部508は前記図16に示した従来の音声合成部508と同様の高調波音源部501と高調波加算部502から構成されており、高調波音源部501は、基本周波数 ω_0 および有声／無声情報に基づいて、該基本周波数 ω_0 および有声とされた周波数バンドに対応するその高調波信号を発生し、前記スペクトル包絡情報 $|B(\omega)|$ に基づいて、それら各周波数の信号の振幅を制御して、高調波加算部502でそれらを加算する。

【0029】また、無声音声合成部509は、対称ラン

ダム系列発生部 201、反対称ランダム系列発生部 202、ランダム系列抽出部 203、逆周波数変換部 204 およびフレーム補間部 205 から構成されている。そして、有声／無声判別情報 $V[k]$ が無声を示す周波数バンドでは、対称ランダム系列発生部 201 で発生される中心対称ランダム系列と反対称ランダム系列発生部 202 で発生される中心反対称ランダム系列を雑音信号の周波数スペクトル系列の実部と虚部と見なし、前記ランダム系列抽出部 203 において前記 2 つのランダム系列から対応する無声の周波数バンドを抽出し、そのパワーを対応した無声のハーモニクス帯域のパワーと同じになる様に振幅調整した後、逆周波数変換部 204 で逆フーリエ変換 (IFFT) することによりその実部を得てこれを無声フレーム信号とし、フレーム補間部 205 で 1 つ前の無声フレーム信号とフレーム間で線形補間することにより無声音声を生成した後、加算器 507 において前記生成した正弦波群と加算して合成音声を得るようにしている。

【0030】すなわち、本発明の音声復号方法においては、従来方式の様に雑音音源から FFT により雑音音源に対応した周波数スペクトルを作成するのではなく、対称ランダム系列発生部 201 および反対称ランダム系列発生部 202 から発生されるランダム雑音シーケンスから、直接、雑音音源に相当する周波数スペクトルを作成する方法を取っている。そして、その周波数スペクトルから無声周波数バンドに対応する周波数帯域を抽出し、逆 FFT によって実時間軸での無声音声を生成した後、フレーム間で補間重みが 1 になる線形補間によって必要なフレーム長さの無声音声を合成するようにしている。これにより、1 回の逆フーリエ変換のみで無声音声を合成することが可能となり、演算量を少なくすることが可能となる。

【0031】ここで、前記ランダムシーケンスの発生に条件を設定する必要がある。これは、逆 FFT によって時間軸シーケンスに変換した場合に、虚数部分が発生せず、FFT スペクトルの全パワーが実時間軸シーケンスに現れるようにする条件と同じである。この条件は式 (10) で表現できる。

$$\begin{aligned} \text{Re}(S_w(m)) &= \text{Re}(S_w(-m)) \\ \text{Im}(S_w(m)) &= -\text{Im}(S_w(-m)) \end{aligned} \quad (10)$$

ここで、 $S_w(m)$ は周波数スペクトルと見なしたランダムシーケンス、 Re は実部、 Im は虚部、 m は周波数スペクトルで $m=0$ の時が DC 成分を表す。

$$A[m] = \sum_{n=0}^{M-1} s(n)w_s(n)e^{-j2\pi mn/M} \quad \text{for } 0 \leq m \leq (M-1) \quad (12)$$

【0037】フレームエネルギー計算部 103 は、周波数スペクトル $A[m]$ から、前記適応窓によるエネルギー

【0032】前記図 1 に示した本発明の音声符号化方法が適用された音声符号化装置についてさらに詳細に説明する。図 1 において、音声入力端子 301 から入力された 8 kHz 程度の標準化周波数で標準化された音声デジタル信号は、基本周波数推定部 401 に入力され、ここで、例えば 20 msec の時間間隔毎に一定長の音声セグメント (フレーム) を取り出し、そのセグメント内での音声基本周波数 ω_0 を推定する。基本周波数の推定方法には、自己相関を用いる方法や、ケプストラムを用いる方法がある事は前述の通りである。

【0033】また、前記音声デジタル信号は、有声強度計算部 404 およびスペクトル包絡計算部 405 にも入力される。スペクトル包絡計算部 405 において、第 2 スペクトル計算部 111 は、該音声セグメントを固定窓処理部 110 でハミング窓等の窓関数で窓処理した信号を高速フーリエ変換 (FFT) することにより離散的な周波数スペクトル値 $B[m]$ を計算する。デジタル音声入力信号の標準化周波数を f_s とし、256 点の FFT を行った場合、計算される周波数スペクトル $B[m]$ は次の式 (11) で表される周波数間隔 f_d 毎に計算される。

$$\text{【数 11】} \quad f_d = \frac{f_s}{256} \quad (\text{Hz}) \quad (11)$$

【0034】スペクトルパワー計算部 112 は、前記基本周波数推定部 401 で推定された基本周波数 ω_0 の整数倍の周波数を中心とし該基本周波数と等しい帯域幅を有する各ハーモニクス帯域毎に、前記周波数スペクトル $B[m]$ の二乗和の平方根を算出し、これをスペクトル包絡値 $|B(\omega)|$ として出力する。

【0035】有声強度計算部 404 は、適応窓処理部 101、第 1 スペクトル計算部 102、フレームエネルギー計算部 103、バンドエネルギー計算部 104、対数変換部 105、バンドハーモニクス振幅計算部 106、バンドハーモニクス幅計算部 107、バンドハーモニクス数計算部 108、有声強度判定部 109 により構成される。

【0036】適応窓処理部 101 は、音声信号 $s(n)$ に対し、前記基本周波数推定部 401 で推定された音声基本周波数 ω_0 から前述した式 (6) ~ (8) で適応的に設定した長さのハミング窓で窓処理を行い、第 1 スペクトル変換部 102 で式 (12) に示す FFT により音声セグメントの周波数スペクトル $A[m]$ を得る。ここで M は FFT サンプル数である。

【数 12】

低下分を補償したフレームの平均エネルギー (「フレームエネルギー」あるいは「フレームパワー」と呼ぶ) E

fを式(13)により計算する。ここで、第2項により窓関数によるエネルギー減少を補償している。

【数13】

$$Ef = \sum_{m=0}^{M-1} |A[m]/M|^2 \cdot \frac{M}{\sum_{n=0}^{M-1} w_s(n)^2} \quad (13)$$

【0038】バンドエネルギー計算部104は、各周波数バンド毎の平均エネルギー（「バンドエネルギー」あるいは「バンドパワー」と呼ぶ） $Eb[k]$ ($k=1, \dots, K$) を計算するものであり、バンドエネルギー $Eb[k]$ は、第kバンドのスペクトル区間を $[ak, bk]$ とすると、次の式(14)で表わされる。

【数14】

$$Eb[k] = \sum_{m=ak}^{bk-1} |A[m]/M|^2 \cdot \frac{M}{\sum_{n=0}^{M-1} w_s(n)^2} \quad (14)$$

$$LA[m] = 20 \log_{10} |A[m]| \quad (16)$$

【0040】バンドハーモニクス振幅計算部106は、各周波数バンド内のハーモニクス振幅 Ah または Bh を計算する。図8を用いて、ハーモニクス振幅の評価方法について説明する。ハーモニクス振幅はスペクトル振幅 $|A[m]|$ のデータ列の極大値とその最近傍の極小値の差であるが、ハーモニクス振幅が線形で表されている場合にはその振幅はスペクトル強度に比例して増減する。そこで、スペクトル振幅の極大値 $H0$ とその前後の極小値 $H1$ 、 $H2$ との差を極大値 $H0$ で正規化した値 $Ah1$ 、 $Ah2$ をハーモニクス振幅の評価値とすれば、スペクトル強度に

関係しないハーモニクス強度が評価できる。ここで、 A

$$Bh1 = -20 \log_{10}(H1/H0) = 20 \log_{10}(H0) - 20 \log_{10}(H1)$$

$$Bh2 = -20 \log_{10}(H2/H0) = 20 \log_{10}(H0) - 20 \log_{10}(H2) \quad (18)$$

$$Bh = \min(Bh1, Bh2)$$

この $Bh1$ や $Bh2$ はハーモニクスのピークからの減衰量をデシベル単位で表したもので、前記図3に示した音声のスペクトル振幅測定結果からも、スペクトル周波数やスペクトル振幅の影響が少ない妥当なハーモニクス強度の評価単位である事がわかる。

【0041】バンドハーモニクス幅計算部107は、前記対数変換部105の出力を受けて、前記各スペクトル振幅極大値の直前の極小値と直後の極小値との間の周波数間隔をそのハーモニクスの幅として算出する。バンドハーモニクス数計算部108は、前記対数変換部105

for ($m = ak + 1; m < bk - 1; m++$) {

if ($20 \log_{10} |A[m]| > \max(20 \log_{10} |A[m+1]|, 20 \log_{10} |A[m-1]|)$)

then (ハーモニクス数 Hn に1を加える。)

ここで、バンドの周波数範囲を基本周波数 ω_0 の3倍に設定する場合には、 ak 、 bk は、

【数15】

$$ak = \text{floor}\left(\frac{M}{2\pi}(3(k-1) + 0.5)\omega_0\right)$$

$$bk = \text{floor}\left(\frac{M}{2\pi}(3k + 0.5)\omega_0\right) \quad (15)$$

になる。ただし、 $\text{floor}(x)$ は x を越えない最大の整数を示す。

【0039】対数変換部105は、前記第1スペクトル計算部102で計算された周波数スペクトル値 $|A[m]|$ の対数値を計算して、対数スペクトル振幅列 $LA[m]$ を計算する。

【数16】

$h1$ と $Ah2$ の小さい方をハーモニクス振幅評価値 Ah とすると、

【数17】

$$Ah1 = (H0 - H1)/H0 = 1 - H1/H0$$

$$Ah2 = (H0 - H2)/H0 = 1 - H2/H0 \quad (17)$$

$$Ah = \min(Ah1, Ah2)$$

となる。または、スペクトル極大値とスペクトル極小値の比でハーモニクス強度を表したハーモニクス評価値 Bh で評価しても良い。すなわち、

【数18】

$$Bh1 = -20 \log_{10}(H1/H0) = 20 \log_{10}(H0) - 20 \log_{10}(H1)$$

$$Bh2 = -20 \log_{10}(H2/H0) = 20 \log_{10}(H0) - 20 \log_{10}(H2) \quad (18)$$

$$Bh = \min(Bh1, Bh2)$$

の出力を受けて、前記式(15)で示した周波数バンドの周波数スペクトル範囲に含まれるハーモニクスの数 Hn を計算する。ハーモニクス数の計算は、FFTで得られる離散的周波数 ak から bk まで周波数スペクトル振幅 $20 \log_{10} |A[m]|$ とその前後のスペクトル振幅 $20 \log_{10} |A[m-1]|$ 、 $20 \log_{10} |A[m+1]|$ を比較し、いずれの値よりも多きければ m 番目のスペクトルはスペクトルの極大点でハーモニクスの中心周波数に最も近いスペクトルであると判断する。すなわち、

【数19】

for ($m = ak + 1; m < bk - 1; m++$) {

if ($20 \log_{10} |A[m]| > \max(20 \log_{10} |A[m+1]|, 20 \log_{10} |A[m-1]|)$)

then (ハーモニクス数 Hn に1を加える。)

(19)

}

【0042】ここで、計算されたスペクトル対数値をそのまま用いて上記方法により極大値の数を数えると、スペクトル雑音の影響を受けて雑音によるスペクトル極大値を数えてしまう弊害があるため、予めスペクトル雑音除去を行い雑音による誤計数を防止するようにしている。このスペクトル雑音除去の方法について図9を参照して説明する。図9の(A)と(C)はスペクトル雑音のある場合を示しており、 $m+1$ と $m+2$ のスペクトル振幅が逆転している。連続した4本のスペクトルの組に対してスペクトル振幅の差分の符号が $-+-$ または $++$ の場合には極大値があり、その極大値はそれぞれ $m+2$ 番目か $m+1$ 番目に現れて、その極大値の振幅は $m+1$ 番目と $m+2$ 番目の振幅の差になることがわかる。そこで、 $m+1$ 番目と $m+2$ 番目のスペクトルの差が雑音レベルを考慮したある閾値より小さければ、 $m+1$ 番目と $m+2$ 番目のスペクトル振幅を両者の平均値に置きかえる事により、図9の(B)と(C)に示す様にスペクトル雑音を除去する事が出来る。

【0043】有声強度判定部109は、前記フレームエネルギー計算部103、バンドエネルギー計算部104、対数変換部105、バンドハーモニクス振幅計算部106、バンドハーモニクス幅計算部107およびバンドハーモニクス数計算部108で算出された、フレームエネルギー(フレームパワー)Ef、バンドエネルギー(バンドパワー)Eb[k]、ハーモニクス振幅Hpw[n][0]、ハーモニクス幅Hpw[n][1]、ハーモニクス数Hnの各パラメータを用いて、バンド毎の有声強度V[k]を計算し出力する。ここで、Hpw[n][0]はその周波数バンドにおけるハーモニクスの振幅(AhあるいはBh)の上位n個までの振幅で、Hpw[n][1]はそれに対応するハーモニクス幅を表している。

【0044】この有声強度V[k]は、入力パラメータを閾値判定して得られる2値の有声/無声の判定結果でも良いし、入力パラメータの判定値の重み付き加算による多値レベルを持った判定結果でも良い。あるいは、入力パラメータの判定値の重み付き加算結果を閾値判定して得られる2値の判定結果であっても良い。有声強度V[k]として2値の判定結果を用いる場合は、各バンド毎に有声か無声かを切り替えて音声合成を行うこととなる。多値の判定結果(例えば、0.0~1.0の範囲の値をと

$$LA[m] = 20 \log_{10} |A[m]|$$

次に、1406でスペクトル雑音除去を行う。このスペクトル雑音除去の処理フロー(ステップ1421~1428)については後述する。

【0047】次に、有声強度V[k]の判定を行う。まず、ステップ1407でフレーム全体のパワー(フレームパワー)Efが所定の閾値Th0より小さいフレームは音声パワーが少なく雑音領域と考えられる場所であるので、ステップ1416ですべてのバンドを無声と設定し

る)の場合には、個々のバンド毎に合成した有声と無声の合成音声を重みつき加算合成して最終合成音声を生成すればよい。

【0045】図10、図11は、図1における前記有声強度計算部404の処理内容を示す処理フロー図である。有声強度計算が開始されると、ステップ1401で基本周波数 ω_0 と周波数スペクトル振幅 $|A[m]|$ を受け取り、1402でそれらをデータ領域に設定する。ここで基本周波数 ω_0 を使用しているが、これはバンド数やバンドの周波数範囲を決定するのに使用するものであり、有声強度の判定に直接使用するものではない。ステップ1403ではバンド数Kを決めるが、各バンドにh本のハーモニクスを含む様に設定した場合には、バンド数Kは、

【数20】

$$K = \text{ceil} \left(\frac{0.5fs/\omega_0}{h} \right) \quad (20)$$

で計算される。ここで、 $\text{ceil}(x)$ はx以上で最小の整数を示す。例えば、 $h=3$ 程度に設計してバンド数Kを計算する。 h と ω_0 が決まれば、前記式(15)により各バンド番号 $k=1, 2, \dots, K$ に対して各バンドの中に入るFFTスペクトルの周波数領域 $[a_k, b_k]$ を計算する。

【0046】ステップ1404では、フレームパワーEf、および、バンドパワーEb[k] ($k=1, 2, \dots, K$)を、前記式(13)、式(14)より計算する。

【数21】

$$Ef = \sum_{m=0}^{M-1} |A[m]/M|^2 \cdot \frac{M}{\sum_{n=0}^{M-1} w_s(n)^2} \quad (13)$$

【数22】

$$Eb[k] = \sum_{m=a_k}^{b_k-1} |A[m]/M|^2 \cdot \frac{M}{\sum_{n=0}^{M-1} w_s(n)^2} \quad (14)$$

次に、ステップ1405でスペクトル振幅 $|A[m]|$ の対数を取りデシベルに変換した対数振幅 $LA[m]$ を計算する。

【数23】

$$(16)$$

でバンドループに入らずに終了する。一方、フレームパワーEfが閾値Th0より大きいフレームに対しては、ステップ1408~1415のバンドループに入る。このバンドループでは、まずステップ1409でその周波数バンドのパワーEb[k]を評価し、所定の閾値Th1以下の場合はそのバンドにはエネルギーが少ないと判断して、無声 $V[k]=0$ と設定する(ステップ1414)。閾値Th1より大きい場合は、ステップ1410でバンドのハーモ

ニクス振幅 $H_{pw}[n][0]$ とハーモニクス幅 $H_{pw}[n][1]$ とハーモニクス数 H_n を計算する。なお、フローチャート中では、ハーモニクス振幅 $H_{pw}[n][0]$ とハーモニクス幅 $H_{pw}[n][1]$ とをまとめて $H_{pw}[n][2]$ と表記している。このステップ 1410 のハーモニクス振幅とハーモニクス幅、ハーモニクス数の計算の処理フロー（ステップ 1430～1450）については、後述する。

【0048】次に、ステップ 1411 でハーモニクス数 H_n を評価し、設定したバンド内ハーモニクス数 h との差がある範囲外（閾値 Th_{20} 以下、閾値 Th_{21} 以上）であれば無声 $V[k]=0$ と判定する（ステップ 1414）。例えば、バンドあたりのハーモニクス数 h を 3 本と設定した場合は 2 以下、4 以上は無声音と判定する。次に、ステップ 1412 でハーモニクス振幅 $H_{pw}[n][0]$ とハーモニクス幅 $H_{pw}[n][1]$ を評価し、それぞれ所定の閾値 Th_3 、 Th_6 より小さい場合はハーモニクス振幅が少ないか、そのバンド幅が狭い無声音と判定する（ステップ 1414）。ハーモニクス幅の閾値は適応窓処理部 101 で設定された窓関数により適応的に設定される。たとえば

（8）式の適応ハミング窓の場合は、ハーモニクス幅は適応ハミング窓スペクトル分布の正、負の第 1 の谷間距離で表されるメインローブのスペクトル幅と関連づけて考えるのが妥当である。ハミング窓のメインローブのスペクトル幅 M_w は窓長 L_w と FFT 段数 M により（21）式で計算されるので、 Th_6 はこの値と関連して実用的な閾値を設定する。

【数 24】

$$M_w = 4 \cdot \frac{M}{L_w} \quad (21)$$

同様に、 Th_3 はハミング窓の第 1 の減衰量に関連しており、基本周波数と適応窓処理部の窓長が前記（6）式の条件を満たしている場合には、ハミング窓の第 1 の谷の減衰量をベースとして実用的な値を設定する。以上で無声音と判定されなかったバンドは、ステップ 1413 で有声バンド（ $V[k]=1$ ）と設定する。以上の動作を各バンド毎に最大 K バンドまで計算し各有声強度 $V[k]$ に設定し終わると、ステップ 1417 でこの有声強度計算部 404 の処理を終える。

【0049】このようにして、フレームパワー E_f について閾値判定し（1407）、各バンドについて、そのバンドパワー $E_b[k]$ について閾値判定し（1409）、ハーモニクス数 H_n について閾値判定し（1411）、さらに、ハーモニクス振幅 $H_{pw}[n][0]$ とハーモニクス幅 $H_{pw}[n][1]$ について閾値判定（1412）して、これらの判定結果から 2 値（0 あるいは 1）の有声強度 $V[k]$ を決定することができる。なお、前述のように、有声強度 $V[k]$ は、このような 2 値の情報に限られることはなく、前記各閾値判定の結果に対してそれぞれ所定の重みを付け、これらを加算することにより、多値（例えば、0.0～1.0 の範囲）の有声強度を算出するようにして

もよい。あるいは、重み付け加算の結果を所定の閾値を用いて判定し、2 値の値とすることもできる。

【0050】次に、前記ステップ 1406 のスペクトル雑音除去のサブルーチン 1421～1428 の処理内容について説明する。ステップ 1421 で受け取ったスペクトル振幅の対数値 $LA[*]$ に対して、ステップ 1422～1427 のノイズ除去ループに入る。このノイズ除去ループでは、連続した 4 点の周波数スペクトル振幅の中に小さな極大点があるかどうかをチェックしている。もし小さな極大点があれば、その極大点に最も振幅値が近いスペクトル振幅との平均を取り、両者のスペクトル振幅をその平均値で置き換え、小さなスペクトル極大点を無くす処理を行う。

【0051】まず、ステップ 1423 で、連続した 4 点の差分 d_1 、 d_2 、 d_3 を計算し、その符号 s_1 、 s_2 、 s_3 を計算する。次に、ステップ 1424 で s_1 と s_3 が同じ符号で s_2 と異なるかを判定する。その結果が真である場合は極大点が真中の 2 点のいずれかである。前記図 9 に示した様に、極大点の振幅は s_1 と s_2 が両方正、両方負の場合いずれでも同じ d_2 の絶対値で表され、ステップ 1425 により $|d_2|$ が所定の閾値 Th_4 より小さな場合には、ステップ 1426 で $LA[m+1]$ と $LA[m+2]$ をそれらの平均値で置きかえる事で小さな極大値の平滑除去を行う。以上の平滑化処理を最後の 4 点のスペクトルが取れるまでバンド内で繰り返し、スペクトル雑音による極大点の除去を行っている。なお、前記図 9 から、極大点を除去すれば、その直前または直後の極小点も同時に取れる事がわかる。

【0052】次に、前記ステップ 1410 のハーモニクス数 H_n とハーモニクス振幅 $H_{pw}[n][0]$ 、ハーモニクス幅 $H_{pw}[n][1]$ の計算サブルーチン 1430～1450 の処理内容を図 11 を用いて説明する。まず、ステップ 1431 で対数スペクトル振幅 $LA[m]$ 、基本周波数 ω_0 、バンド番号 k （ $k=1, 2, \dots, K$ ）、バンドスペクトル範囲 $[a_k, b_k]$ を入力として処理を開始する。ステップ 1432 で、極大値の数を計数する極大値数カウンタ N_{pk} 、極小値の数を計数する極小値数カウンタ N_{btm} 、極大値の振幅を格納する極大値メモリ $A_{pk}[*]$ 、極小値の振幅を格納する極小値メモリ $A_{btm}[*]$ 、ハーモニクスの振幅を格納するハーモニクス振幅メモリ $H_{pw}[*][0]$ 、ハーモニクスの帯域幅を格納するハーモニクス幅メモリ $H_{pw}[*][1]$ 、ハーモニクスの数を計数するハーモニクス数カウンタ H_n をそれぞれ 0 に初期化する。また、ハーモニクス幅の開始点 mb_1 と終了点 mb_2 をそのバンドのスペクトル開始点 a_k に設定する。

【0053】次に、ステップ 1433 でピーク・ボトム計算ループ（ステップ 1433～1448）に入り、ステップ 1434 で対数スペクトル振幅 $LA[m]$ が $LA[m-1]$ 、 $LA[m+1]$ より大きい場合は、 $LA[m]$ が極大値と判定しステップ 1435 へ移動する。ステップ 1435

で、発見された極大値がバンド内で始めて発見された場合であるかを検出し、始めて検出された場合には、ステップ1436で極大値数カウンタNpk及び極小値数カウンタNbtmに1を設定し、その極大値LA[m]を極大値メモリApk[1]に、初期値LA[ak]を極小値メモリAbtm[1]に記録する。始めて検出されたものでないときは、ステップ1437で極大値数カウンタNpkをインクリメントし、極大値LA[m]を極大値メモリApk[Npk]に記録する。

【0054】一方、前記ステップ1434のピーク検出でピークでないと判定された場合には、引き続いてステップ1438で極小値であるかの判定を行う。この判定は、前記ステップ1434の極大値判定と同様な手法で行い、この結果極小値と判定された場合には、ステップ1439で極小値数カウンタNbtmをインクリメントし、極小値LA[m]を極小値メモリAbtm[Nbtm]に記録する。さらに、ハーモニクス幅の計算のため、mb1をmb2に更新し、mb2には現在のスペクトル周波数mを設定する。極大値、極小値判定ともNoと判定された場合は、ステップ1441でボトム／ピーク検出ループの最後であるかを判定し、最後のループの場合はステップ1442に進み、極大値数カウンタ値Npkと極小値数カウンタ値Nbtmが同じであるか否かを判定する。同じである場合には、ステップ1440で極小値数カウンタNbtmをインクリメントし、極小値メモリAbtm[Nbtm]にLA[bk]を記録し、ハーモニクス幅の計算のため、mb1をmb2に更新し、mb2には現在のバンドの最終スペクトル周波数bkを設定する。この手順ですべての極大値が検出され、その前後の極小値も記録される。

【0055】次に、ステップ1443で、極小値が検出された時点でその前に極大値があるかを判定し、もしあれば、その極大値を新たなハーモニクスとしてステップ1444でその振幅Haを計算する。ステップ1444では、その極大値と前後の極小値との振幅差の平均値をハーモニクス振幅Haとしている。しかし、ハーモニクス振幅形状の対称性を重要と考えて判定する場合には、前記式(18)で示した様に、最小値でHaを計算しても良い。次に、ステップ1450でハーモニクス幅Hwを計算し、ステップ1445で、Haを所定の閾値Th5と比較し、閾値より大きい場合だけ、ハーモニクス数Hnを更新し(ステップ1446)、上位n個のハーモニクス振幅をHpw[n][0]にハーモニクス幅Hpw[n][1]を記録する(ステップ1447)。ステップ1447のmaxN(Hpw[n], Ha, Hw)は、HaがHpw[n][0]の配列要素の最小値より大きい場合にハーモニクス振幅を示す第1の配列要素の最小値と置き換え、同時にハーモニクス幅を示すその配列番号の第2要素をHwと置き換える関数を示している。すべてのピーク／ボトム計算ループを終えると、ステップ1449で、バンド内でのハーモニクスの数Hnと上位n個のハーモニクス振幅と幅Hpw[n][2]を

戻している。以上、有声強度計算部404の処理内容を詳細なフロー図で説明した。

【0056】次に、前記図2に示した本発明の音声復号方法が適用された音声復号装置における無声音声合成部509について詳細に説明する。前述のように対称ランダム系列発生部201は中心対称ランダム系列を発生し、反対称ランダム系列発生部202は中心反対称ランダム系列を発生する。ここで、中心対称ランダム系列は、系列中のある1点(中心とする)からみて、振幅極性ともに左右対称(すなわち、中心で折り返したときに、中心の左右にある系列が完全に一致している状態)であるランダム系列をいい、中心反対称ランダム系列は、中心からみて振幅は左右対称であるが極性は反転しているランダム系列をいう。実際には、逆周波数変換部204において実行される逆フーリエ変換処理の段数(逆FFT段数とよぶ)の1/2の長さのランダム系列を発生させ、これを発生順序の逆方向に複写することにより、前記中心対称のランダム系列を発生させることができ、また、前記逆FFT段数の1/2のランダム系列を発生させ、これを発生順序の逆方向に極性を反転して複写することにより、前記中心反対称のランダム系列を発生させることができる。

【0057】このようにして前記対称ランダム系列発生部201および反対称ランダム系列発生部202で発生された2つのランダム系列は、ランダム系列抽出部203に供給され、ここで、該2つのランダム系列を周波数スペクトル系列の実部と虚部と見なして、前記有声／無声音情報により無声と指定された周波数バンドに対応する区間の系列が抽出されるとともに、抽出したスペクトルパワーを前記スペクトル包絡情報B(|ω)|に対応した無声ハーモニクス帯域のパワーと同じになる様に振幅調整される。この振幅調整された無声ハーモニクス帯域スペクトルは逆周波数変換部204において逆フーリエ変換されて時間領域の信号に変換され、無声フレーム信号に対応する該逆FFTの段数と同じ数の時間軸データ系列が得られる。

【0058】このようにして得られた例えば256個(逆FFT段数が256段の場合)のデータは、フレーム補間部205に入力され、前記音声セグメントの更新周期に対応したデータ数(例えば、20msec周期であれば160個のデータ)に補間合成される。これは、前のセグメントから得られた時間軸データとこのセグメントの時間軸データとを補間重みの和が1になる条件で線形補間するものである。このようにして合成された無声音声は、前記加算器507に供給され、前述した有声音声合成部508からの有声音声と加算される。

【0059】図12は上述した無声音声合成の処理フローを示す図である。まず、ステップ1602でパラメータ復号化部からハーモニクスのスペクトル包絡情報|B(ω)|、音声基本周波数ωo、バンド有声無声音情報V[k]

を受け取り、バンド数Kmaxを(20)式により再生する。ここで、各バンドに含まれるハーモニクス数hはシステムで予め決定されている。フレームサイズFsizeは予め設定している音声セグメント更新間隔で、fs=8000Hz、セグメント更新周期を10msecとした場合はFsize=80である。ステップ1603はIFFT段数Mに256を使用する場合はFFTスペクトル数の実部、虚部それぞれ256要素を0に初期化する。ステップ1604はランダムFFTスペクトル発生の初期化で、システム立ち上げ時のみ必要で、連続音声再生時には不要である。

【0060】ステップ1605～1614は処理フレー

$$u(u+1) = 171 \cdot u(n) + 11213 - 53125 \cdot \text{floor}\left(\frac{171 \cdot u(n) + 11213}{53125}\right) \quad (22)$$

【0061】ステップ1607～1613はハーモニクスループであり、各バンドに含まれるハーモニクス数の回数だけ処理される。まず、ステップ1608で、各バンド内にある第1次ハーモニクスのスペクトル範囲[a1, b1]を(23)式により計算し、ステップ1609で、その範囲だけ前記ランダムシーケンスu[n]から抜き取り、u[a1, b1]を取り出す。

【数26】

$$\begin{aligned} a1 &= \text{floor}\left(\frac{M}{2\pi}(l - 0.5)\omega_0\right) \\ b1 &= \text{floor}\left(\frac{M}{2\pi}(l + 0.5)\omega_0\right) \end{aligned} \quad (23)$$

ここで、Mは逆FFT(IFFT)の段数である。

【0062】次に、ステップ1610で取り出した抽出

$$U2(m) = |B(\omega)| \cdot U1(m) \cdot M$$

ここで最後の項のMは、ステップ1616でのM段IFFTの出力が実時間信号レベルに合うために必要な係数である。次に、ステップ1612でレベル調整された抽出スペクトルを対応するFFTスペクトルバッファS_real[M], S_imag[M]に設定する。

【0063】以上の処理を各バンド、各バンド内のハーモニクス毎に実行した後、ステップ1615に進み、前記式(10)の関係を満足する負の周波数のFFTスペクトル部分を設定し、ステップ1616でのM段のIFFTで得る時間軸信号の実部に、全スペクトルエネルギーが集まり、虚部には信号が現れない様になっている。ス

$$ws(n) = \begin{cases} 0 & ; (|n| > L2) \\ 1.0 & ; (|n| < L1) \\ 1 - (n - L1)/(L2 - L1) & ; (L1 \leq |n| \leq L2) \end{cases}$$

$$\begin{aligned} uv(n) &= u_real_prev(M/2 + n) \cdot ws(n) \\ &+ u_real_now(M/2 - (Fsize - n)) \cdot ws(-(Fsize - n)) \end{aligned} \quad (26)$$

【0064】なお、以上においては、音声符号化方式としてIMBE方式を採用した音声符号化伝送装置の音声

ムのバンド数の回数だけ処理されるループで、無声バンドの周波数帯域の無声音のスペクトルをハーモニクス帯域毎に再生加算し、フレームの無声音スペクトル全体を再生するループである。ステップ1606は要素数がIFFT段数の半分のランダムシーケンスを順次生成するもので、例えば、前記IMBE方式では(22)式で発生させており、本方式でも同様の方式で生成しても良い。但し、ここでは実部と虚部用の2系列のランダムシーケンスu[n]を発生し、u[n]は直流成分を除去するため、53125/2を減じた値を用いる。

【数25】

スペクトルのパワーが1になる様に(24)式によりスペクトルを正規化する。ここでU(m)は実部と虚部のランダムシーケンスu_real[m]とu_imag[m]をベクトルで表現したものであり、U1(m)は正規化された抽出スペクトルu_Real[m]とu_Imag[m]をベクトル表現したものである。

【数27】

$$U1(m) = U(m) / \sqrt{\sum_{m=a1}^{M-1} |U(m)|^2} \quad (24)$$

ステップ1611はハーモニクスのスペクトル包絡情報|B(ω)|により、ハーモニクス帯域内のエネルギーが元音声の帯域内エネルギーと同じになる様に(25)式により抽出スペクトルを振幅調整する。

【数28】

$$(25)$$

ステップ1617では現在フレームと前フレームから得たそれぞれサンプル数Mの信号から、式(26)に示す補間関数ws(n)のフレーム間補間によりフレームサンプル数(Fsize)の無声音の復号音声を得る。そして、ステップ1618で無声音の復号音声を報告し、図2の加算部507にて別に合成された有声音合成部508の音声と加算し最終の合成音声を得る。図13は、前記フレーム補間関数ws(n)の例を示す図である。ここで、L1は補間関数の一定レベル範囲、L2は補間最大範囲、L1～L2間は直線補間範囲である。

【数29】

符号化パラメータ抽出部に本発明の音声符号化パラメータの取得方法を適用した場合を例にとって説明したが、

本発明の音声符号化パラメータ抽出方法および装置は、これに限られることなく、MELP (Mixed Excitation Linear Prediction) 方式など、1 フレームの周波数スペクトルを複数の周波数バンドに分割し、各周波数バンド毎に有声/無声を判定する場合に全く同様に適用することができる。また、音声復号方式としてIMBE方式を採用した音声符号化伝送装置の無声音の音声復号部に適用した場合を例にとりて説明したが、本発明の無声音復号方法および装置は、これに限られることなく、MELP方式など、1 フレームの周波数スペクトルを複数の周波数バンドに分割し、各周波数バンド毎に有声/無声を判定する場合にも同様に適用することができる。

【0065】

【発明の効果】以上述べた様に、本発明の音声符号化パラメータの取得方法および装置によれば、音声基本周波数により、音声ハーモニクス高調波スペクトルがお互いに分離する様に適応的な可変長窓処理を行った音声セグメントから周波数スペクトルを求め、検出するハーモニクス振幅とハーモニクス幅とハーモニクス数の信頼性を高め、さらに、音声セグメントのパワー、音声セグメントを複数の周波数バンドに分割した各周波数バンドのパワーにより、有声強度あるいは有声/無声音情報を取得しているため、音声基本周波数の変化によらず、ハーモニクスのレベルの低い部分へのハーモニクス高調波雑音の影響が少ない有声強度判定を行うことが可能となる。したがって、スペクトル雑音に対し、誤り耐性の強い音声符号化パラメータの取得方法を提供することができる。また、本発明の音声復号方法および装置によれば、無声音の復号において、ランダム雑音からFFTによってランダム周波数スペクトルを作成せず、直接ランダム周波数スペクトルを生成する事が出来るため、無声音信号の復号時にFFTとIFFTの計算のうち、IFFT計算のみで無声音を合成する事が出来、従来の方法に比べて演算負荷の小さな音声復号方法を提供する事が出来る。

【図面の簡単な説明】

【図1】 本発明の音声パラメータ取得方法が適用された音声符号化パラメータ抽出部の構成を示すブロック図である。

【図2】 本発明の音声復号方法が適用された音声合成部の構成を示すブロック図である。

【図3】 音声セグメントの対数スペクトル振幅（有声部）の例を示す図である。

【図4】 音声セグメントの対数スペクトル振幅（無声

部）の例を示す図である。

【図5】 ハーモニクスのスペクトル形状を説明するための図である。

【図6】 第1の周波数分析窓長の設定例を示す図である。

【図7】 第1の周波数分析窓の形状の例を示す図である。

【図8】 ハーモニクス振幅を説明するための図である。

10 【図9】 スペクトル雑音除去を説明するための図である。

【図10】 有声強度計算処理の流れを示すフローチャートである。

【図11】 ハーモニクス計算処理の流れを示すフローチャートである。

【図12】 無声音合成処理の流れを示すフローチャートである。

【図13】 無声音のフレーム間補間を説明するための図である。

20 【図14】 音声符号化伝送装置の構成を説明するための図である。

【図15】 従来の音声符号化パラメータ抽出部のブロック図である。

【図16】 従来の音声合成部のブロック図である。

【図17】 正規化スペクトル誤差とピッチ周波数誤差の関係を説明するための図である。

【符号の説明】

101 適応窓処理部

102 第1スペクトル計算部

30 103 フレームエネルギー計算部

104 バンドエネルギー計算部

105 対数変換部

106 バンドハーモニクス振幅計算部

107 バンドハーモニクス幅計算部

108 バンドハーモニクス数計算部

109 有声強度判定部

110 固定窓処理部

111 第2スペクトル計算部

112 スペクトル包絡計算部

40 201 対称ランダム系列発生部

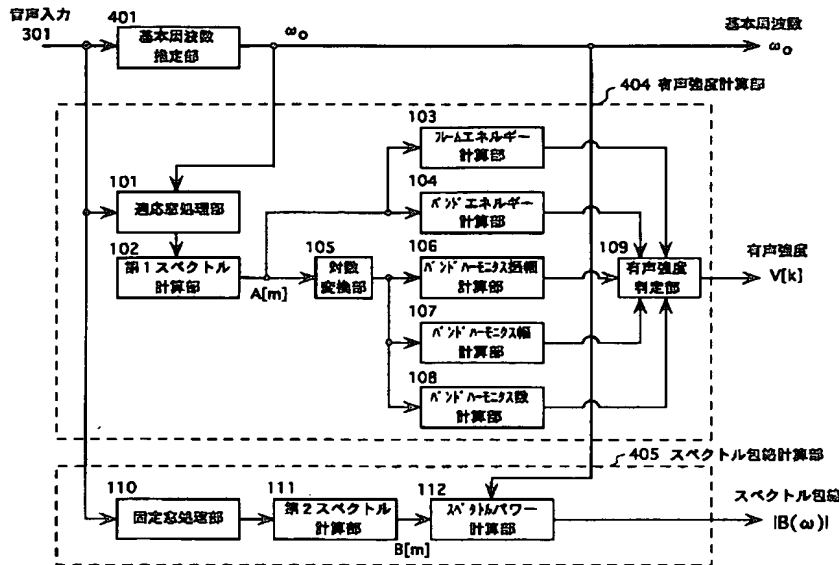
202 反対称ランダム系列発生部

203 ランダム系列抽出部

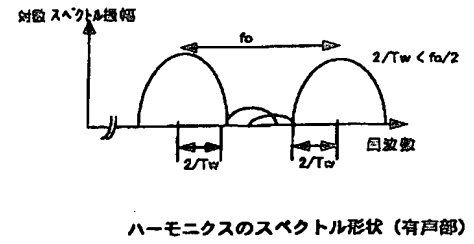
204 逆周波数変換部

205 フレーム補間部

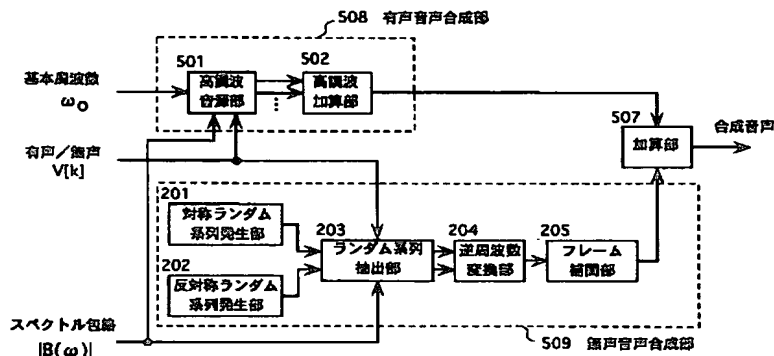
【図 1】



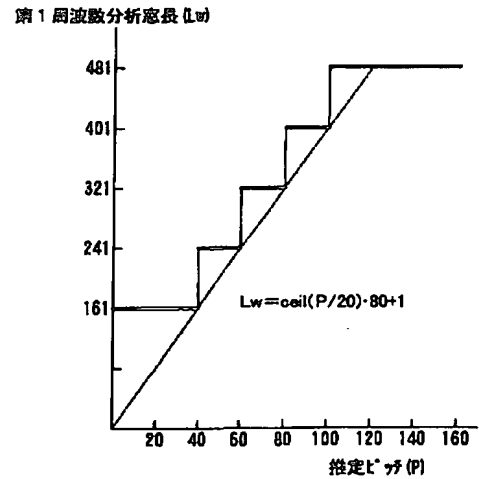
【図 5】



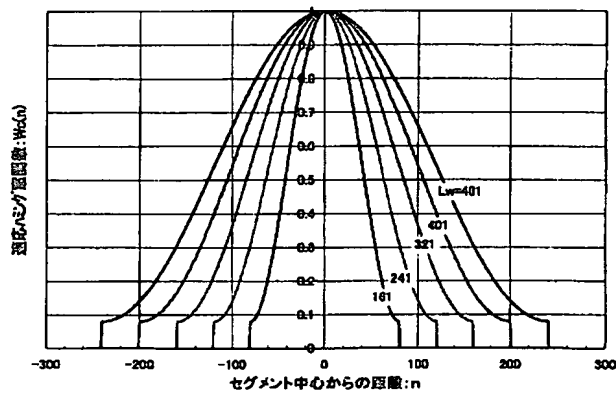
【図 2】



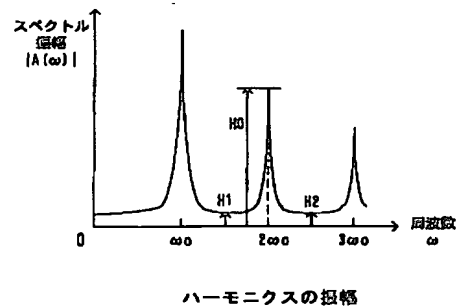
【図 6】



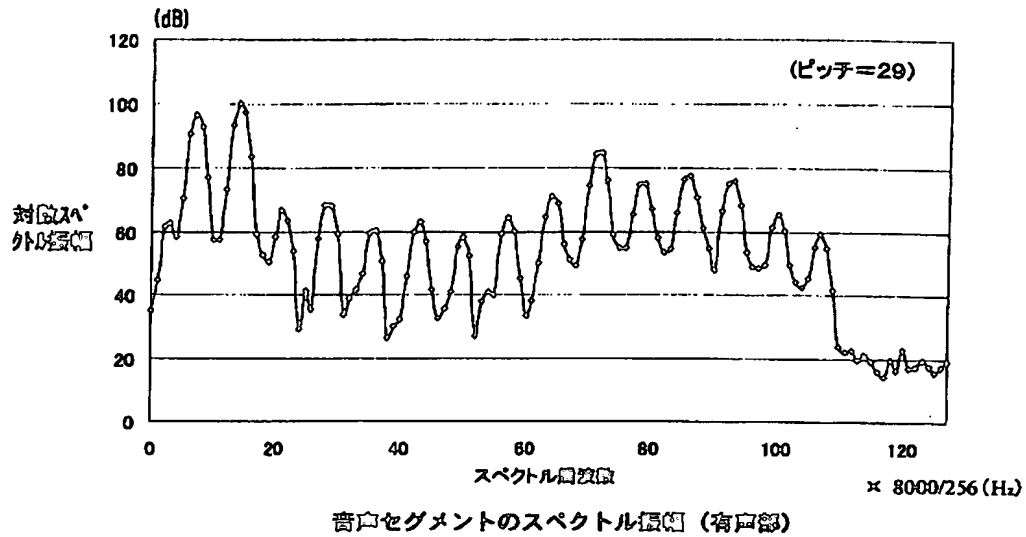
【図 7】



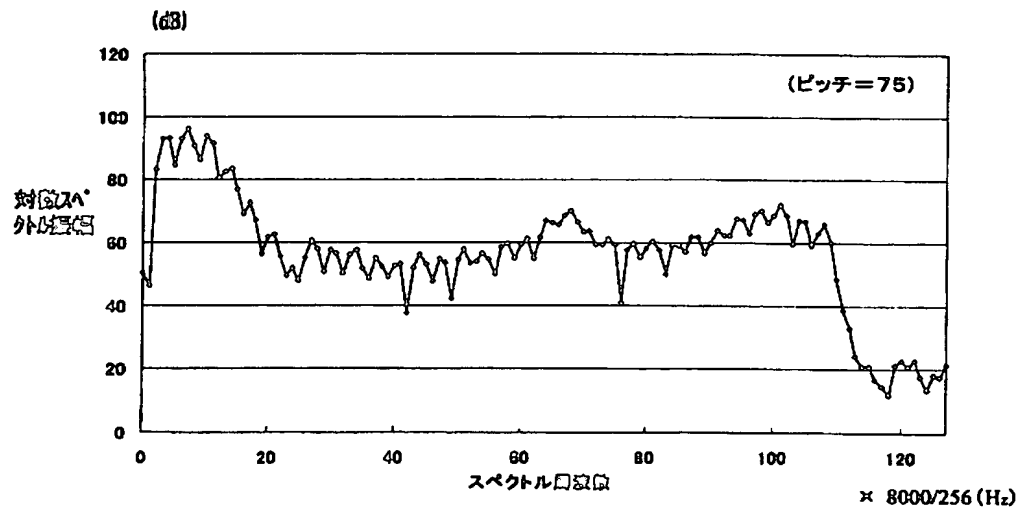
【図 8】



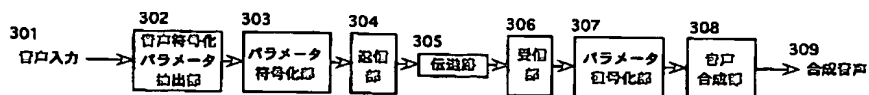
【図 3】



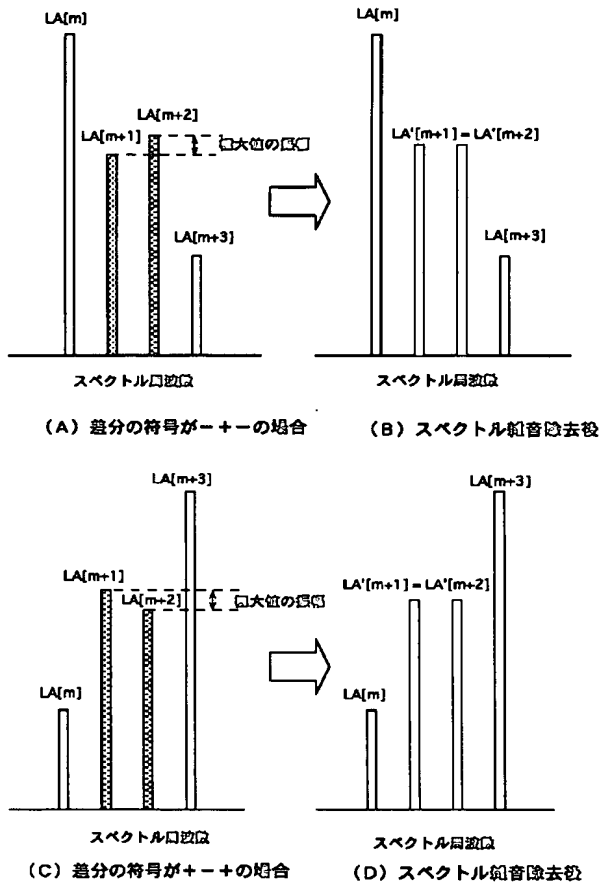
【図 4】



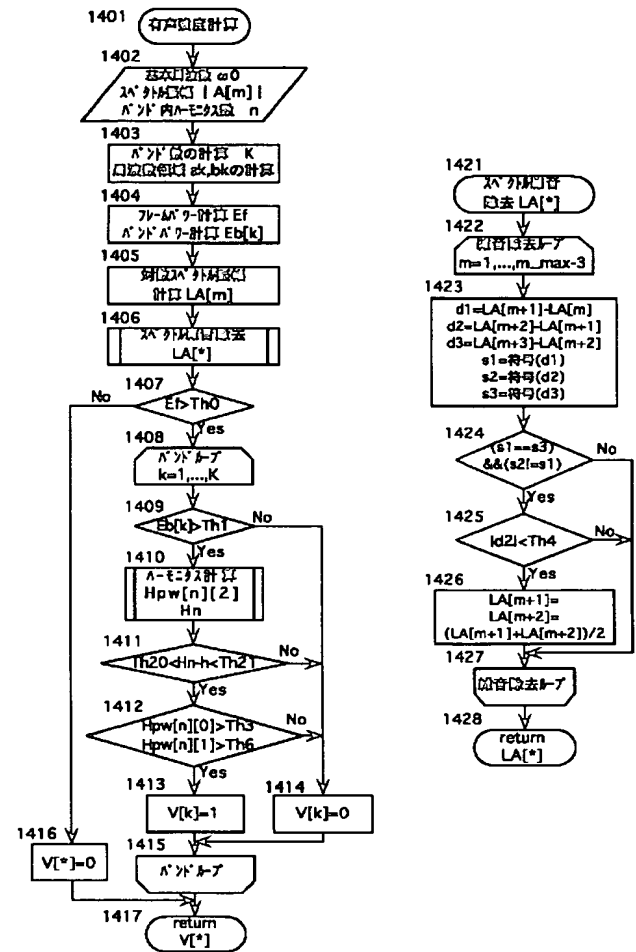
【図 14】



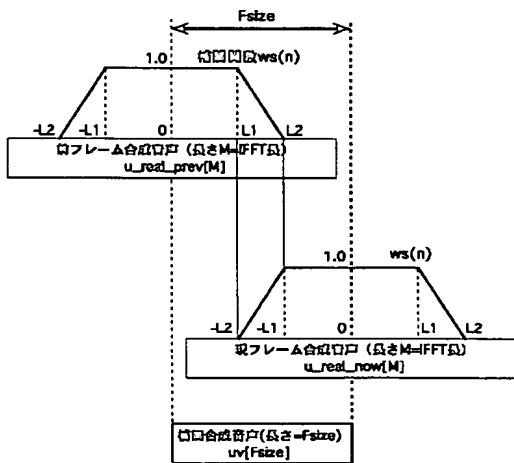
【図 9】



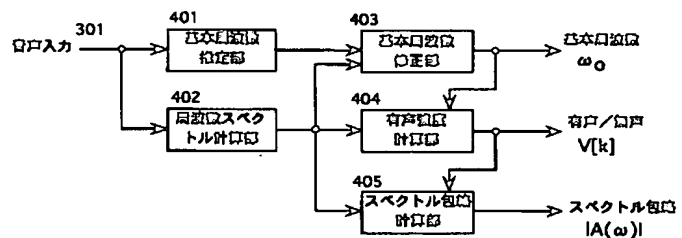
【図 10】



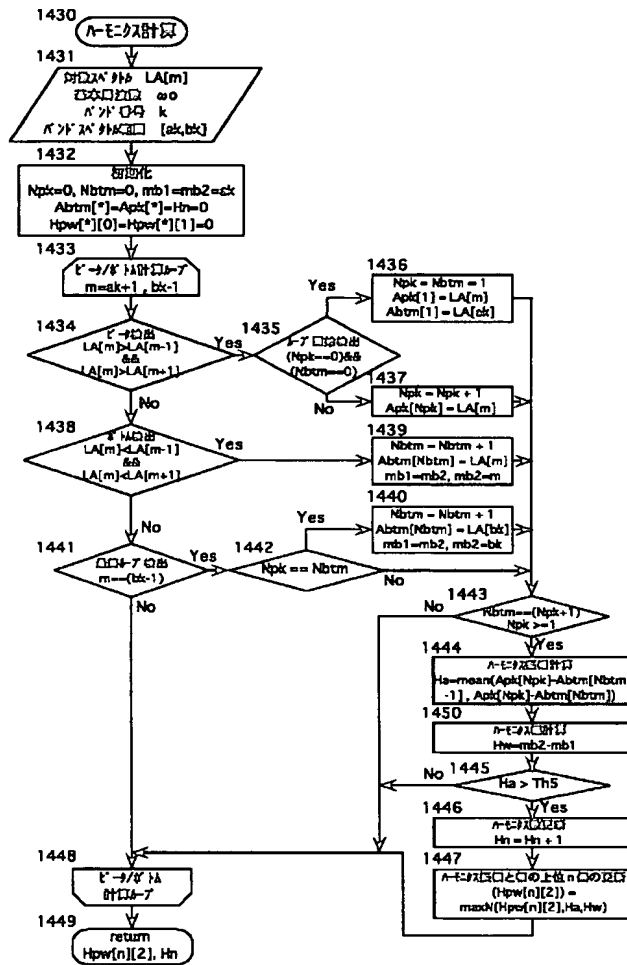
【図 13】



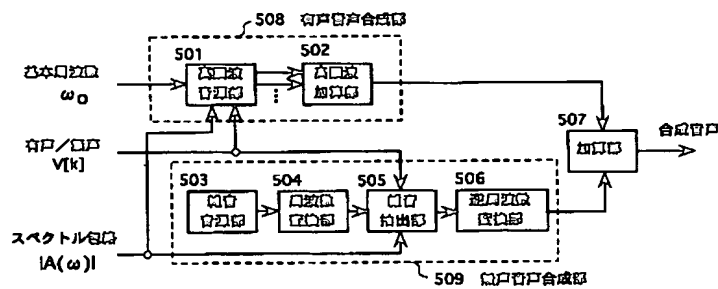
【図 15】



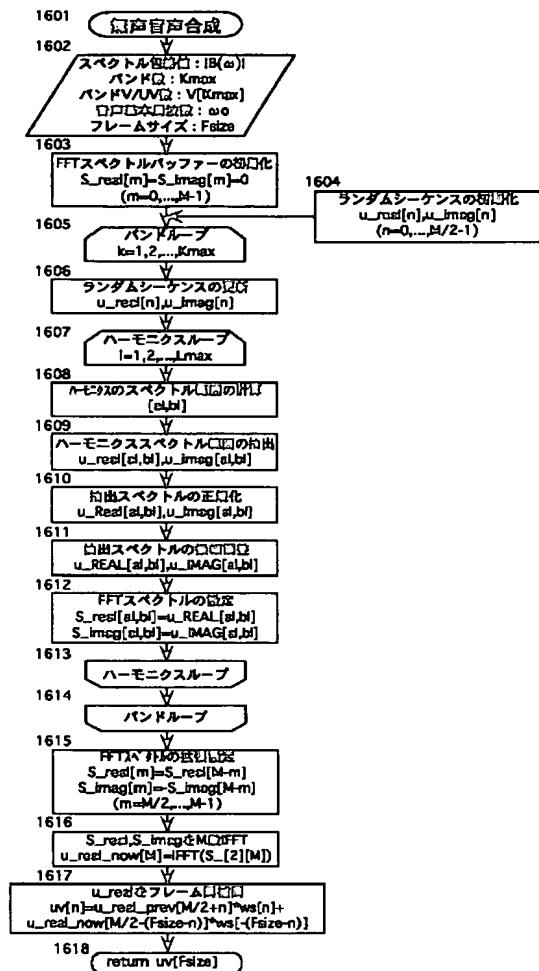
【図11】



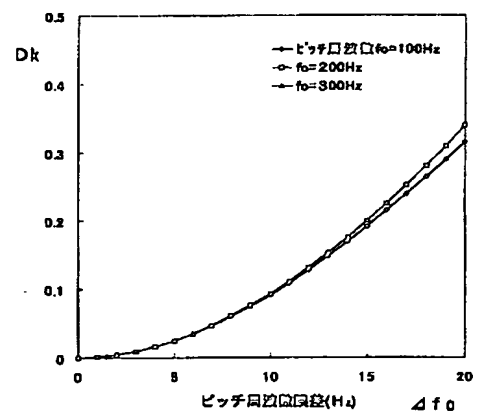
【図16】



【図12】



【図17】



正規化スペクトル包絡

フロントページの続き

(72) 発明者 佐々木誠司

神奈川県横須賀市光の丘 3 番 2 号 株式会
社ワイ・アール・ピー高機能移動体通信研
究所内

F ターム(参考) 5J064 AA00 BB04 BB12 BC08 BC14
BC18 BD01
5K041 AA00 BB08 CC01 DD01 EE23
HH01 HH10 HH24 HH44
9A001 EE04 HH15